DOCUMENT RESUME

ED 081 803 TM 003 164

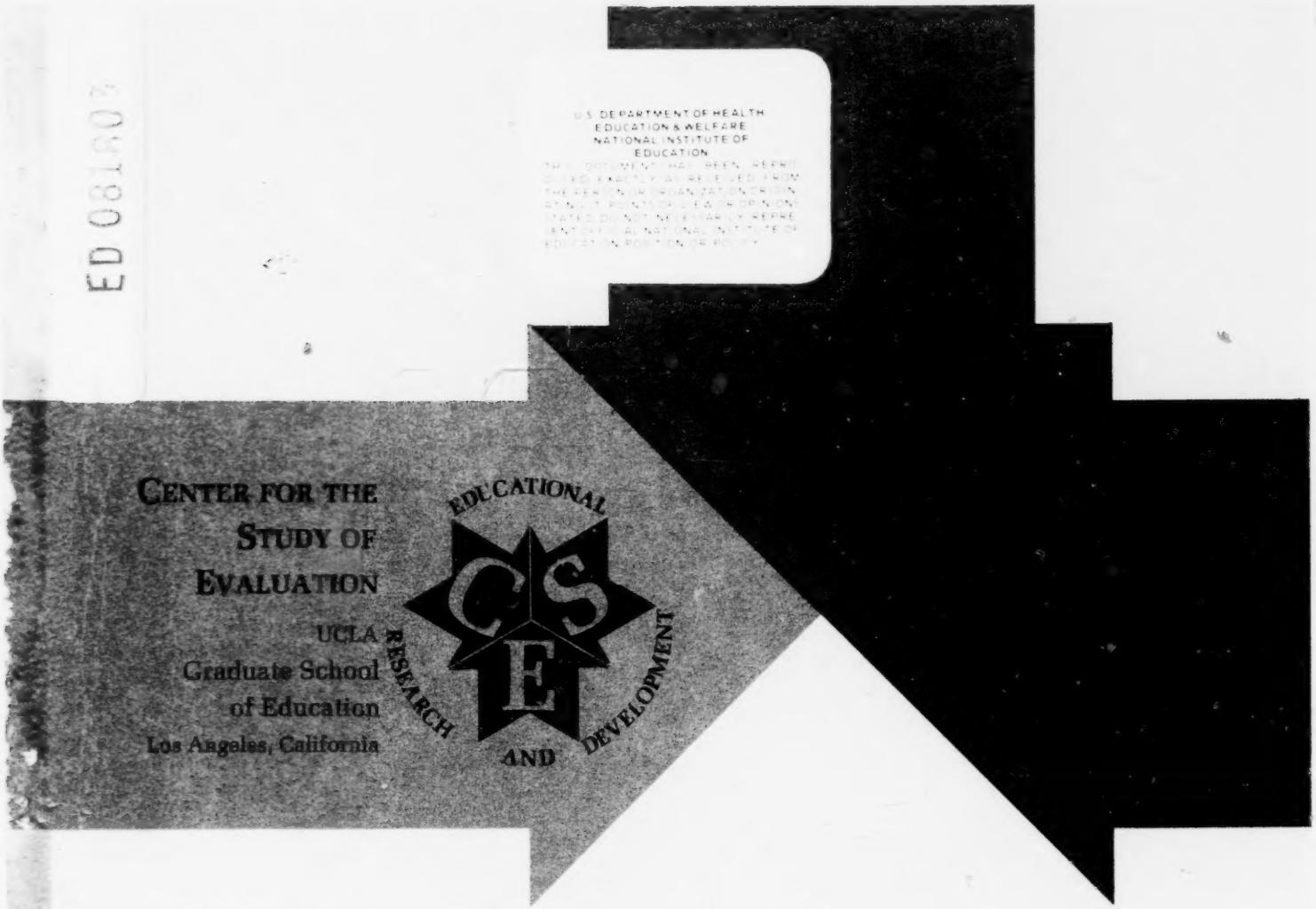
AUTHOR	Hoepfner, Ralph
TITLE	The Validity of Tests of Social Intelligence.
INSTITUTION	California Univ., Los Angeles. Center for the Study of Evaluation.
SPONS AGENCY	Office of Education (DHEW), Washington, D.C.
REPORT NO	CSE-R-85
PUB DATE	Apr 73
NOTE	35p.; Address presented to the German Congress of Psychology (Saarbrucken, Germany, October 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	Behavioral Science Research; *Evaluation Techniques; Factor Analysis; *Intelligence Tests; Models; Performance Tests; Predictive Validity; Psychological Tests; *Social Psychology; Speeches; *Test Construction; Tests; *Test Validity
IDENTIFIERS	*Social Intelligence

ABSTRACT

The rationale underlying tests cf social intelligence and some of the problems inherent in those tests are discussed. To measure social intelligence, paper and pencil tests were developed which were "situation free." These tests employed "stereotypic behavior of individual others." The stimuli used in the behavioral tests were photographs of people, artists' drawings, and cartoons. The types of responses were generally limited to marking on an answer sheet or writing a response. A series of tests for behavioral creativity that did not use paper and pencil, answer sheets, or writing were developed. Four types of performance tests were developed and administered to a sample of 30 people along with the paper-and-pencil tests. A new set of tests, approximately 25 to 30, was developed for behavioral production. Conclusions reached as a result of two validation studies are that people with high verbal intelligence don's need behavioral intelligence to perform well on the tests, and individuals with lower intelligence to perform well on the tests, and individuals with lower intelligence sometimes still achieved very high behavioral scores. Studies have also been done relating social intelligence to certain demographic variables. A study investigating the prediction of academic achievement correlated the behavioral-cognition tests with grades at a junior college; correlations of about .50 were found with achievement in various course grades. Three studies of the use of behavioral-cognition tests for predicting vocational success are discussed. Portions of the tests relating to factors of behavioral intelligence and factors of social creativity and their tests are provided. (DB)

ED 081807

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION



THE VALIDITY OF TESTS OF SOCIAL INTELLIGENCE

Ralph Hoepfner

CSE Report No. 85
April 1973

FILMED FROM BEST AVAILABLE COPY

CENTER FOR THE
STUDY OF
EVALUATION

MARVIN C. ALKIN
DIRECTOR

UCLA Graduate School of Education

The CENTER FOR THE STUDY OF EVALUATION is one of eight educational research and development centers sponsored by the U.S. Office of Education, Department of Health, Education and Welfare. Established at UCLA in 1966, under provisions of the Cooperative Research Act, CSE is devoted exclusively to the area of evaluation.

The mission of the Center is to conduct research and development activities for the production of new materials, practices and knowledge leading to the development of systems for evaluating education which can be adopted and implemented by educational agencies. The scope of activities includes the development of procedures and methodologies needed in the practical conduct of evaluation studies of various types, and the development of generalizable theories and concepts of evaluation relevant to different levels of education.

This publication is one of many produced by the Center toward its goals. Information on CSE and its publications may be obtained by writing to:

Dissemination Office
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California 90024

ED 081803

TM 003 164

THE VALIDITY OF TESTS OF SOCIAL INTELLIGENCE*

Ralph Hoepfner

CSE Report No. 85
April 1973

Evaluation Technologies Program
Center for the Study of Evaluation
UCLA Graduate School of Education
Los Angeles, California

*Invited address presented to the German Congress of Psychology, Saarbrucken, Germany, October, 1972.

Much of the research reported herein was supported by the U.S. Office of Education and the Office of Naval Research through the Aptitudes Research Project, University of Southern California, J.P. Guilford, Director.

Historical Background

The very early historical philosophers, most certainly the early authors, poets, and dramatists, recognized the fact that human beings communicate with each other in a way that is not verbal. That is, we can sense other peoples' intentions; we can sense their feelings; we can understand what their desires are without being told in words.

It remained until about 1920 that this kind of intelligence was given a name. Thorndike was speaking to teachers at the time, trying to help the teachers to understand the many different varieties of thinking skills that they could teach to their students. He stated that, on one hand, we have concrete intellegence--how to manipulate things, to know about things. In another way, we have abstract intelligence--knowledge of words or numbers. But Thorndike was not satisfied that the two types of intelligence exhausted all the ways that one could conceive of intelligence, so he added another concept which he called "social intelligence."

For some time, however, nothing much was done scientifically about social intelligence. There was, of course, an early test of social intelligence called the George Washington Test of Social Intelligence. Psychologists and educators demonstrated on many occasions that the attempts to measure social intelligence by the George Washington scale were not effective: the test scores correlated far too highly with scores of verbal tests or with scores of numerical tests.

Theoretical Inclusion in the SI Model

When Guilford formulated the structure of intellect model between 1956 and 1960 (Guilford, 1959), he already had three types of intelligence of which he was quite confident: figural, symbolic, and semantic. At that time,

he and many other psychologists had demonstrated many of the factors comprising these types of intelligence. For example, Thurstone, Spearman, and Vernon had demonstrated verbal factors and number factors. But Guilford went out on a limb and included in his model of intelligence a whole new dimension which he called social intelligence or behavioral intelligence. Guilford's model of intelligence hypothesized thirty separate and distinct abilities all having to do with social intelligence. The model specifies five operations (what you can do with behavioral information or how you process it) and six products (the results or the kind of thinking element that goes on in the intellect). Figure 1 illustrates the behavioral "slice" of Guilford's Structure-of-Intellect Model.

Breadth and Nature of Social Intelligence

We should make a clear distinction about what Guilford had in mind when he talked about social intelligence. This distinction will arise again and again in our discussion. Guilford's model is concerned with intellectual skills or cognitive skills, so we must differentiate these cognitive skills or aptitudes from other things. What other things could one confuse with cognitive skills?

One thing that could be confused is affect or personality traits. That is; what I know, what I sense you are communicating to me, must be quite different from the affect that I impose upon that communication. We are not concerned with whether I am friendly to you or hostile to you or whether I am shy or whether I am neurotic or psychotic. Instead, we are concerned with how my intellect processes the behavioral information that you communicate to me and that I communicate to you.

We also must avoid confusion with another element that is commonly associated with social intelligence, and that I call knowledge. We all have certain

Figure 1

The Behavioral "Slice" of Guilford's Structure-of-Intellect Model

	Units	Classes	Relations	Systems	Trans-formations	Impli-cations
Cognition	CBU	CBC	CBR	CBS	CBT	CBI
Memory	MBU	MBC	MBR	MBS	MBT	MBI
Divergent Production	DBU	DBC	DBR	DBS	DBT	DBI
Convergent Production	NBU	NBC	NBR	NBS	NBT	NBI
Evaluation	EBU	EBC	EBR	EBS	EBT	EBI

social knowledge, the "Emily Post" type of knowledge. One can demonstrate social skills by how one behaves. If one drinks coffee, does one lift the saucer? Where does one put one's fork? Does one stand when a lady enters the room? Does he offer his hand to a lady? Such "cultured" behaviors exemplify a type of formally learned social knowledge, but it is not social intelligence as defined by Guilford's model, because those skills can be transmitted purely verbally. I suspect that is how most of us learn such skills. We are told that when a lady enters a room, a gentleman stands for her. When your professor enters the room, perhaps, you stand for him. You show respect, but it is not behavioral; it is verbal and formal. You have formally learned it. You've been told to do it.

Another type of social knowledge is stereotypic group knowledge. One might know certain things about groups of people (usually we know wrong things about people when we know stereotypes). If I were to say to you "gypsy," you might have an image based upon your knowledge of what gypsies are. They are wild; they're great lovers perhaps; they cheat; they steal (at least in America that's what we "know" about gypsies). But this may or may not be true--it is a stereotype. Guilford's model does not address stereotype knowledge either. It is not behavioral knowledge; it is, once again, verbal knowledge. While I may never have met a gypsy, I have certain verbal knowledge about them.

The Measurement of Social Intelligence

Given the limitations above, the problems that had to be solved at the University of Southern California were concerned with how to measure aptitudes of behavioral perception and production. Purely for the purposes of conserving time and money, we adopted paper-and-pencil tests for our measurement strategy. Any other kind of test employed for factor analytic purposes,

where one needs to have about 200 examinees (any kind of interpersonal interaction observation schedule or videotape interaction record or social interaction that must be done person-to-person), would be far too expensive, first to administer and second to score. So it was decided instead to use paper-and-pencil tests. Sample items of many tests appear as an appendix to this paper.

The Nature of the Paper-and-Pencil Tests

There are several problems associated with the effective use of paper-and-pencil tests, however. The greatest problem revolves around the situations or contexts in which behavior occurs. What I mean is that the same behavior in different situations means different things. If I were to raise my hand to you now, it would mean something quite different than if I were very angry with you, pushed you against the table and raised my hand to you. The different situation has made the behavioral expression of raising my hand mean something quite different. With paper-and-pencil tests, one cannot take the time to tell what the situation is. That is, one cannot say "here is a picture of a man with his hand raised" and then explain what has happened so that the examinee knows the situation. Instead, we developed tests which we would call "situation free." When one attempts to develop a test that is situation free, one begins to lose much of what is important in social stimulation because the situation does have a very strong determining effect upon what one perceives.

We instead decided to employ what we would call "stereotypic behavior of individual others." We began looking for types of behavior that are common to most people in the culture, so that if any person exhibits the behavior, it would be understood in the culture. If you were to see my face twisted in extreme anger, whether it is a picture, or whether it is in a movie, or almost anywhere you might see my face that way, you would understand that I was very angry; I was in a rage at something that perhaps you had said or that had

happened. But it is stereotypic behavior you would understand. (We will see shortly whether or not such stereotypic behavior is culturally determined.)

The stimuli that were used in the behavioral tests were printed on paper and were of a number of different types. For example, many of the stimuli are photographs of people. We deliberately set up the behavioral situations that were supervised by a movie director from Hollywood. The movie director took many pictures of many people under very carefully staged situations to get stereotypic emotional expressions. We also had artists who drew facial expressions, postures, parts of bodies, silhouettes.

We also used cartoons. We realized, of course, that many cartoons that one sees in the newspapers are very behavioral. Most cartoons happen to be verbal. The joke, the humor in a cartoon, is of a verbal sort. But many are very behavioral. One laughs at them because he understands that this is the way that he sometimes behaves or this is the way that someone that he knows frequently behaves. We also attempted to use tape recordings of vocal stimuli. The vocal stimuli were of the way one speaks a particular expression. For example, we might present the stimuli (sentence) two times. Although it is the same sentence, I am sure that you will understand the difference. First: "You have to go <u>now</u>?" Second: "You have <u>to go</u> now." In the first instance, I am very hurt that you're about to leave me so early. But not so in the second instance. In the second instance, I am telling you to go. So the way I raise or lower my voice also has behavioral information which is added over and above the semantic meaning of the same words that I have used.

We attempted to have one of each kind of test for each behavioral factor predicted by the SI Model. We were very concerned in our experimental design

that we should not have a factor which would be essentially a photograph factor, a cartoon factor, a silhouette factor, or a vocal tape-recorded factor. Therefore, we counter-balanced the stimuli over hypothesized factors in an attempt to meet a good experimental design.

The types of responses that examinees were to produce were also important to us. We generally limited the types of responses to marking on an answer sheet or to writing a response on the paper with a pencil or with a pen (there were few exceptions to this case).

The problem, of course, occurs that when one is engaging in social intelligence which, keep in mind, is situation free, one is dealing with stereotypic information; it is stereotyped by the culture and maybe by many other cultures. The problem arises of how to score the answers--how do we know which answer is the correct one. We cannot do what we would do with any other type of test. If I were to ask you to write a verbal comprehension test, a vocabulary test, it might present a key word with four other words. You could tell me which word was correct. If I didn't believe you, you could take me to a dictionary, open it to the correct page and say: "Look. Here it is. This is correct." If I still seriously resist, I think perhaps you would consider institutionalizing me, because something is wrong with me if I don't at least partially agree with the standard meaning of the word.

Likewise, if I were to create a test of spatial ability, one of visualizing the rotation of blocks, I would give you a picture of one block among four others, and say, "Mark one of the four blocks that is the same as the given block." If I then say that the answer is this one and you say, no, it is not, I could take a piece of wood, saw off the block, turn it for you, and show you that it is the correct response. And I could <u>prove</u> (or at least convincingly demonstrate) to you that it is correct. In both cases, for a verbal test and

for a spatial test, we have what we call criteria for correctness, an ultimate criterion to which (almost) everyone agrees. Everyone accepts what they see, and most everyone believes dictionaries.

When we consider social intelligence, however, there is a great deal of disagreement on the criteria for correctness. If I were to present to you a test item with a given face and four alternative faces, and I said that one is the correct face, you could disagree. Then if you said to me, "Prove to me that this is the correct face," I have no dictionary, I have no piece of wood to show you that it is correct. So how do I demonstrate that I have a correct answer? The ways that the project used to find the correct answers depended very closely on the idea of cultural stereotype of the stimuli presented. If there is a cultural stereotype, then the majority of people within that culture will respond similarly to the same stimuli. We would say there is a "consensual correctness."

And that is precisely how we went about scoring all of our tests. We developed the tests, we pretested them, and we looked at the responses. We did item analyses of every item and every alternative to find out whether the alternative that we thought was correct was the alternative that most of the examinees chose. We went further, of course, to say, "Is that the alternative that most of the high scoring examinees selected?" If so, then we were quite happy with the item. If people did not select the alternative that we thought was correct, then one of two possibilities ensued.

We, of course, had to look at the item to see what went wrong. If you've ever developed a test, or done an item analysis, you know what generally happens when you look at the results of an item analysis. You look at the results and say, "Oh, of course, why didn't I see that in the beginning, how dumb I am." Then you change the item. But sometimes we didn't see what was

"dumb," and we attempted those items again. Unfortunately, if we still didn't understand what was going wrong with the item, we just discarded it and didn't look at it anymore.

At this point, we might look briefly at the notion of "pan-culturalism" of the responses to our social intelligence tests. The question posed by pan-culturalism is whether the stimuli, the facial expressions that we had used or the body expressions, are stereotypic for a given culture or society at a given time, or whether they are appropriate for all societies at all times. There are many psychologists, and there are many theories, that propose that behavioral intelligence is a genetic thing--that human beings learn or have instinctually a certain vocabulary of behavior intelligence. They would say, for example, that apes have facial expressions. When you see pictures of them, they show their teeth and they make faces, and those faces communicate. So psychologists who believe that there are pan-cultural expressions of behavioral information quote this kind of information and conclude that all people understand certain common behavioral expressions.

A study addressing the pan-cultural issue was done about four years ago by Ekmann, Sorenson, and Friesen (1969), using five different populations. They employed three populations of people in the United States, Brazil, and Japan which they called "literate" cultures. They also utilized two additional cultures--one in New Guinea and one in Borneo that they called "pre-literate" cultures. They developed a test for the recognition of behavioral expressions by using seven different expressions--expressions such as happiness, surprise, fear, anger, disgust, contempt, and sadness (the kinds of emotional responses that perhaps you associate with Schlosberg who studied dimensions of emotional expressions). They chose faces that were uniquely and only expressive of those behaviors and presented the pictures to hundreds of people in each of the cultures.

What they found was that the literate cultures scored very very highly regardless of whether they were in the United States or Brazil or Japan--all scores were high. In the pre-literate cultures the scores were still high--not quite as high, but nonetheless high. So what Ekmann and his colleagues concluded was that the comprehension of emotional expression is pan-cultural, that is, it is genetically determined and inherited from the stock of the human race. It therefore is the universal language.

My contention is that their conclusion was a bit unwarranted. I suspect that the basic emotions--rage, happiness, sadness--are indeed pan-cultural, but those are not always the most important behavioral information that people communicate. They are basic, but not always important. My suspicion is that many of the stimuli that have been used in Los Angeles were culturally specific, that is, they are specific to the United States; they're specific to our culture and they are also specific to our time. If you were to use the tests in Germany and found, for example, that the items did not work well, that you did not get reliable inter-correlations among the items, or that the scores did not exhibit a good distribution, that perhaps shouldn't be too surprising. One could simply state that the behaviors used in the U.S. are not common to the German people. I think, however, that such is not the case. The behaviors we have used are in many ways common to German people too. Not all of them, but many of them.

Tests Utilizing Other Media

In our attempts to measure social intelligence we also tried to use other kinds of tests. Another type of test that we employed yielded some very interesting results in a study in which we were concerned with behavioral divergent production (to refresh your memory, divergent production is not perceiving or recognizing, but producing creatively). If we follow Guilford's model of

intelligence, then there must be behavioral creativity. If I were to give you a behavior to communicate to me, you could do it in many different ways. We would call that behavioral creativity. Presumably great actors and actresses have much of this, and I suspect that successful salesmen have a great deal of it too, because they have to communicate to you to find out what it is you want from them and then try to convince you that what they have is what you want. That must take behavioral creativity.

We developed a series of tests for behavioral creativity that did not use paper and pencil, answer sheets, or writing. In two of the tests, people photographed themselves in response to a given situation. In one instance the situation was, "Your doorbell rings, you open the door, and in front of you stands a friend you haven't seen in years." You say, "My goodness." How would you look when you said "My goodness"? You might make many different faces. You might say "My goodness! (how good to see you)" or "My goodness! (you again?)" There are many different ways of showing surprise that your friend is at the door.

We had two tests similar to that, though they were somewhat different from each other. We also had two tests that were tape recorded. The examinee in this case sat before a tape recorder and verbalized his response. We might give the examinee this situation: "You are at a circus with a friend; you look up at the top of the tent and you say 'Look what's happening.'" The examinee then tape records "Look what's happening," in many different ways, each one communicating a different kind of behavior. It is obvious that one could say "Look what's happening" with interest, with humor, or with terror (perhaps someone was falling from the tent or perhaps the tent was falling down). One could say it many different ways.

Four types of performance tests were developed (photographed and tape recorded) and administered to a sample of thirty people along with the paper-

and-pencil tests of behavioral creativity. We presumed that the responses to those tests would exemplify divergent-production-of-behavioral-units (DBU) or perhaps classes (DBC). We therefore included tests of DBU in the paper-and-pencil battery. We factor analyzed the correlations among the eight variables and got three factors--a paper-and-pencil factor, a vocal factor, and a facial factor. That is not what we hypothesized. We had thought that our paper-and-pencil tests would be general indicators of social creativity and would reflect all kinds of behavioral intelligence, but that was not the case.

Once again the problem arises that, if we are concerned with how well the test of behavioral intelligence relates to how people behave, we are concerned with two possible contaminating influences. The first, of course, is affect. If, for example, I were to correlate scores from a test on social intelligence with video-tape observations of how we behave in a classroom, or how you interact with a client who has come for therapy, or how you interact with your mother or with your children, there is affect involved. Much of what we do is not "intelligent" at all; it is emotional (and many times we regret what we do because we do it in emotion). So the test of social intelligence may not correlate with actual behavior because of the emotional component.

But even if we could partial out that emotional component through some statistical manipulation, we still might not get high correlation between test scores and behaviors. The reason would be the same reason that underlay the findings of the study I related to you. Professor Guilford presumed that the reason the various types of tests did not emerge as one factor was because of what he called "executive functions" (Guilford, 1972). That is, what you can do is not always what you know how to do. Some people cannot raise their

voices, they simply do not have very expressive voices. Some people don't have very expressive faces. I understand from talking with my colleagues here (in Germany) that Americans are good at making many different kinds of faces and we tend to make many of them, perhaps appearing a bit funny at times. While we do make many faces, some people may not be able to. So Guilford hypothesized that not only does one have to know what to do, but one also has to be able to do it. One has to get the muscles of the face working or the larynx of the throat to work correctly, in order to communicate different behavioral states.

The Aptitudes Research Project Tests of Social Intelligence

Very briefly, I should mention that we have completed three factor analyses using the tests I've just described with large samples. The first was the study by O'Sullivan, Guilford, and DeMille (1965) wherein six factors of social cognition were found. Those cognition tests were the tests with which you are probably familiar. We demonstrated six factors that we interpreted to support Guilford's model and to say that Guilford's model is a permissible model. It does explain things and we can support it with data to a large degree. Our primary concerns with that first study were to demonstrate that the behavioral abilities were not merely verbal abilities. Therefore, for each behavioral ability that we demonstrated, we attempted to demonstrate the parallel verbal ability. We demonstrated the existence of each one of the six behavioral factors and that they were distinct from their parallel semantic factors.

We replicated this study where once again the six social intelligence factors were demonstrated (Tenopyr, Guilford, & Hoepfner, 1966), this time not against the verbal factors but against a different set of factors. We were successful at demonstrating the separateness of the social intelligence factors from their parallel symbolic factors, but that did not come as a surprise to us. Anyone who has done a factor analysis using symbolic factors

(involving tests with numbers and letters) knows that those factors will emerge clearly. The symbolic factors showed no confusion with behavioral factors.

In the third major study (Hendricks, Guilford, & Hoepfner, 1969), we were not concerned so much with the behavioral cognition factors, but were concerned with divergent-production factors. We developed a new set of tests, approximately 25 to 30 different new tests that were for behavioral production. We had to limit the tests to the paper-and-pencil variety. Therefore, the examinee was asked to write many responses. We might utilize a picture of a person with an exaggerated expression and then instruct the examinees to: "Write all the things this person might be saying to you." One could write many different things -- perhaps, "I hate you," "I'm going to kill you," or "Why did you do this to me?" The real problem was in scoring the tests, because we had to determine if the response really was a behavioral response and if each response was behaviorally different from each other response. Responses such as "I hate you" and "I will kill you" are really not behaviorally different; they are anger directed at the person. So we would not score both of those responses, we could score only one.

There were many tests like that, wherein examinees wrote responses to pictures, cartoons, verbally described situations, and tape recorded messages. Six factors were demonstrated for the six divergent-production factors. We successfully separated each divergent-production factor from its parallel behavioral-cognition factor and also from its parallel semantic-divergent-production factor.

Studies of Construct Validity

Now we should look at what has happened so far with the tests for the behavioral factors. How have people used them? Have they proved productive in the research that has been done in the United States? I might say that

not too many studies have been done. Many of Guilford's ideas I think are perhaps more popular in places other than America. They do not always meet with a great deal of approval among Guilford's American colleagues.

In a study by Tenopyr (1967) using the behavioral-cognition tests, she predicted English grades and history grades among high school students and their scores on a standardized achievement test. She found correlations ranging between .15 to .30, not very high correlations. But more important than the low level of the correlations was the fact that when she then added standard academic-achievement predictors to those tests and predicted grades, using miltiple regression to find out whether or not social intelligence tests would add to the multiple prediction, she found that they did not. When achievement tests are included, they completely dominate the prediction picture. One could reasonably conclude that, if you wish to predict academic grades in high school, don't use behavioral-intelligence tests. But I don't think that I would need to give you that advice. It makes good sense, I believe, that you would not use those kinds of tests as academic predictors.

In another type of validation study (Hoepfner & O'Sullivan, 1968), we again looked at the relationship between verbal IQ (Henmon-Nelson) and tests of behavioral cognition. We looked not so much at the correlation but at the scatter plots of the scores underlying the correlations. We were interested in how the scores on both types of measures were distributed. What we found was that high-IQ people tended to get high social-intelligence scores; low-IQ people tended to get either low or high social-intelligence scores. What can we conclude from this finding which held true in many cases?

I think there are two important conclusions that can be made. First, people with high verbal intelligence don't need behavioral intelligence to perform well on our tests. They can work our tests verbally. If they are

smart enough verbally, they can out-smart the test constructors and answer correctly because we have not developed good enough controls in our tests to force the examinee to work in one (behavioral) and only one way. As I have stated, we use stereotypes in the tests. A person taking the test can say "This person is angry," "This person is morose," "This person is defiant," "This person is loving." The words "defiant" and "angry" are close in verbal meaning, so he chooses them as similar. What he has chosen is words, not behavior. So we see that the tests are susceptible to high scores simply on a verbal level if one has high enough verbal intelligence. Second, however, and probably more important, is the fact that individuals with lower intelligence (and some subjects in our sample had IQs in the 80s) sometimes still achieved very high behavioral scores. The important conclusion to deduce from this finding was that it is not necessary to have high verbal intelligence to do well in social skills. It is not difficult to think of people who get along very well with others, who manipulate other people, yet who are not very smart verbally. One doesn't think of them as being intelligent, but they do understand and manipulate people. They get by very well in the world providing they don't have verbal or numerical problems to solve. We might go further and say that it would be very wise to find out who those people are and to use them. For example, teachers who have very high IQs are probably not as effective as teachers with lower IQs for children of 3, 4, or 5 years of age. They cannot go down to the level of the children to interact with them. The good teacher would be the kind who will understand the children and whom the children will like. She will nurture in them good feelings for learning. She can't really teach them a great deal, but then we don't formally teach children very much at age 3 or age 4.

We have also done some studies relating social intelligence to certain demographic variables. O'Sullivan (1965) included in her factor analysis a

common index of socio-economic status and found a small but significant relationship that the higher the socio-economic status, the higher the behavioral cognition. One can hypothesize that what underlies that finding is that children with higher socio-economic status have more experience with many different people. They travel more; they meet more people; more people come to their homes. They just have much more behavioral experience and therefore they do much better on behavioral tests. Another hypothesis would relate socio-economic status to verbal IQ, and then view verbal IQ as a mediator in performance on the social intelligence tests. Alternatively one might hypothesize that the stereotypic situations selected for the test items are more common to the higher socio-economic students, and that it is the tests that are biased.

Another study did not use one of Guilford's tests, but used a different test developed along the same lines. The new test of social intelligence (Heussenstamm & Hoepfner, 1970) controlled for a number of things. The test had an equal number of items for each sex, that is, boy items, girl items, man items, woman items; and for two different ages, young children and adults; and also for the four major racial groups in the United States which we would call Caucasian, Negro, Mexican-American, and Oriental. The test items represented the four racial groups, two age groups, and the two sex groups, for a total of 16 different types of items that were counterbalanced in the test and then administered to approximately 200 people in each category. That is, 200 young, oriental girls, aged 7-13, were administered the test; 200 adult, Negro men, were administered the test; and likewise in the same manner for each of the 16 groups, 200 people in each group. Examinees were carefully selected who were approximately in the middle socio-economic status for their group. A number of very complicated hypotheses were generated. Among others, it was hypothesized that each group would score highest on its own items and

maybe also with items of the opposite sex (it was felt that while it's very important to know what people of your own sex are communicating to you, it's probably far more important to know what the other sex is communicating to you, at least if you want to have a happy "adjusted" life). While there were many such hypotheses, not one of them was confirmed. We found only that adults scored higher than children, on the adult's test and the children's test, and that there was no systematic variation for total score or for individual item-race scores.

In a third study looking at demographic characteristics, Shanley, Walker, and Foley (1971) administered a number of the behavioral-cognition tests to children of ages 12, 15, and 18 years to determine whether there was development over those three age levels. They found very significant development using the Guilford tests, which utilize primarily adult stimuli. They found score increases from 12 to 15 to 18 just as one would expect with an academic achievement test. According to the authors, the finding indicates that behavioral intelligence has a developmental sequence just like other aspects of intelligence. From that we may imply, of course, that much of what we know and much of what we talk about as behavioral intelligence is learned.

Studies of Predictive Validity

In a study investigating the prediction of academic achievement, Allen (unpublished) correlated the behavioral-cognition tests with grades at a junior college. He found once again correlations of about .50, which were very high correlations, with achievement in various course grades. But once again he did what Tenopyr had done. He used those tests in a multiple regression of the grades along with verbal tests and he found the same thing that Tenopyr found--that verbal tests contribute all the predictive variance in a multiple-regression sense. Of course, that's not a surprise because the

criterion that is being predicted is not behavioral. You do not get good grades in junior college primarily by being socially intelligent. You get good grades primarily by being verbally or numerically intelligent. There are exceptions of some people who know how to manipulate their professors by being very very socially intelligent.

The last subject I would like to discuss is the use of the behavioral-cognition tests for predicting not academic achievement but vocational success. Three studies are relevant here. In the personnel department of IBM (Gershon, unpublished), there is the problem of selecting managers for various computer-related activities. It seems reasonable that the behavioral intelligence tests should be predictive of how well a manager can deal with people and how well he can handle the complex interpersonal relationships with which he must work. On this assumption the tests were used to predict the employment success of managers. The success of the managers was measured primarily by supervisor ratings and by length of time on the job.

In a very similar study, the same set of tests were used with probation officers. When a criminal is released from jail, he usually must report to someone once a week or once a month to show that he is behaving and that he is a better person. Of course if he's behaviorally intelligent he knows how to give the probation officer what the probation officer wants to hear. O'Sullivan (unpublished) used the behavioral-cognition tests to study the success of probation officers, success being measured by their supervisor's ratings. In both the Gershon study with IBM managers and the O'Sullivan study with probation officers, there is a great deal of logic in using the behavioral-cognition tests. Because the type of work that those people do is largely behavioral, the tests should correlate with success criteria, and they did. The correlations sometimes went as high as .60, and very few of them were

less than .40. But along with the good news came some bad news. When verbal tests were added to the multiple regression equations, the same thing happened as happened before. The verbal tests took all the predictive variance and the behavioral tests made no significant contribution. While it was disappointing to learn those results, we can see why they occurred. The reason, if I may be allowed to give you my hypothesis, is due to the nature of the tests. As you may recall, the tests are not situation-specific; a probation officer's chores are very situational and so is a manager's chores, so that it may have been unfair to use these very situationally generalized tests as predictors.

Perhaps a more specific, situational test should have been developed. In the case of the probation officer, who deals with people who don't behave well according to society's standards, the test should perhaps be composed of items written with that kind of situation in the background: Is the person lying to you, is he cheating on you, is he telling you the truth? In other words, instead of being situation free (but stereotypically specific) as many of these tests are, the tests might better have been situation specific.

In the most recent study of a vocational prediction, O'Sullivan (unpublished) used the social intelligence tests to predict the success of nurses. It was quite logically hypothesized that much of what nurses do must has a social-intelligence component. That is, not only do nurses have to take temperatures and give medicine, but they also must reassure the patients, help them to get well, create a better climate for them psychologically. Therefore, it follows that they should be socially intelligent to be successful. When the success criterion was the ratings of both supervisor and peers, there were no meaningful correlations at all between tests and criteria. The correlations were approximately zero, but the hypothesized reason for the finding was that in the hospital in which the study was performed the hierarchical structure in the nursing staff was very strong. The head nurse was

right; even when she was wrong, she was right. The hierarchy was so strong that every nurse rated the nurse above her as being good and every nurse rated the nurse below her as being bad; so, of course, whether the tests worked or not really is unknown. The criterion didn't work, and we had no way of telling how useful the tests might be.

In summary, it was my intent to present to you the rationale underlying the tests of social intelligence and to expose some of the problems inherent in those tests. These problems may well explain the reluctance of many psychologists to utilize the tests in practical and research situations. It is hoped, however, that the modest beginnings will lead to improvements in the test instruments and subsequently to the increased utilization of assessment of one of our important human assets, social intelligence.

References

Ekmann, P., Sorenson, E.R., & Friesen, W.V. Pan-cultural elements in facial displays of emotion. Science, 1969, 164, 86-88.

Guilford, J.P. Executive functions and a model of behavior. Journal of General Psychology, 1972, 86, 279-287.

Guilford, J.P. Personality. New York: McGraw-Hill, 1959.

Hendricks, M., Guilford, J.P., & Hoepfner, R. Measuring creative social intelligence. Reports from the Psychological Laboratory, No. 42. Los Angeles: University of Southern California, 1969.

Heussenstamm, F.K., & Hoepfner, R. Inter-Person Perception Test. Hollywood, California: Monitor, 1970.

Hoepfner, R., & O'Sullivan, S. Social intelligence and IQ. Educational and Psychological Measurement, 1968, 28, 339-344.

O'Sullivan, M., Guilford, J.P., & DeMille, R. The measurement of social intelligence. Reports from the Psychological Laboratory, No. 34, Los Angeles: University of Southern California, 1965.

Shanley, L.A., Walker, R.E., & Foley, J.M. Social intelligence: A concept in search of data. Psychological Reports, 1971, 29, 1123-1132.

Tenopyr, M.L. Social intelligence and academic success. Educational and Psychological Measurement, 1967, 27, 961-965.

Tenopyr, M.L., Guilford, J.P., & Hoepfner, R. A factor analysis of symbolic memory abilities. Reports from the Psychological Laboratory, No. 38. Los Angeles: University of Southern California, 1966.

Thorndike, E.L. Intelligence and its uses. Harper's Magazine, 1920, 140, 227-235.

Factors of Behavioral Intelligence

CBU - Cognition of Behavioral Units

The ability to understand the facial and postural expressions of other people.

Expressions CBU01A. Which alternative expresses the same thought, feeling, or intention as the given?

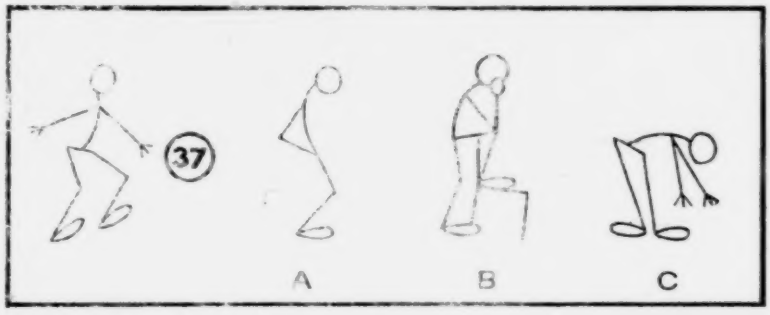
Answer: 4. Score: number of items right plus one-fourth of the number omitted.
Parts: 2; items per part: 18/18, 14/18; working time: 10 minutes.

Faces CBU02A . Which man's face expresses the same feeling or intention as the woman's?

Answer: 4. Score: number of items right plus one-fourth of the number omitted.
Parts: 2; items per part: 11/15, 14/15 working time: 8 minutes.

Stick Figure Expressions - CBU05A. Choose one of three stick-figures that expresses the same feeling or intention as the given (left) figure.

Sample Item:



Answer: C.

Score: Number of right responses minus one-half number wrong.

Parts: 2; items per part: 15/18, 15/18; working time: 18 minutes.

CBC - Cognition of Behavioral Classes

The ability to see similarity of behavioral information in different expressional modes.

Expression Grouping CBC04A. Which alternative expression belongs with the given group of expressions?

Answer: 1. Score: number of items right plus one-fourth of the number omitted.
Parts: 2; items per part: 12/15, 14/15; working time: 10 minutes.

Picture Exclusion CBC05A. Which photographed expression does not belong with the other three?

(12)

Answer: 3. Score: number of items right plus one-fourth of the number omitted.
Parts: 2, items per part: 10/15, 11/15; working time: 10 minutes.

CBR - Cognition of Behavioral Relations

The ability to understand diadic interactions between other people.

Silhouette Relations CBR05A . Which photograph expresses the individual's feeling or intention in the silhouette relationship? In part one of the test, the alternative pictures are of men. In part two, they are of women.

Answer: 1. Score: number of items right plus one-third of the number omitted.
Parts: 2; items per part: 11/15, 12/15; working time: 10 minutes.

Social Relations CBR02A . Which statement expresses the feeling of the face indicated by the arrow, taking into account the relationship between the faces?

1) I didn't like that movie very much.
2) What a [illegible]re!
3) Who does he think he is, anyway?

Answer: 3. Score: number of items right plus one-third of the number omitted.
Parts: 3, items per part: 5/7, 7/7, 7/7; working time: 6 minutes.

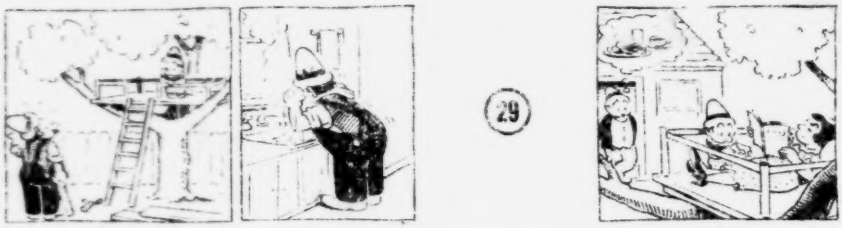
CBS - Cognition of Behavioral Systems

The ability to understand complex interactions among several other people.

Missing Pictures CBS04A. Which alternative completes the story, making sense of the thoughts and feelings of the actors?

Answer: 3.
Score: number of items right plus one-third of the number omitted.
Parts: 2; items per part: 11/14, 10/14; working time: 16 minutes.

Missing Cartoons CBS01A. Which alternative completes the cartoon strip, making sense of the thoughts and feelings of the characters?



Answer: 4. Score: number of items right plus one-fourth of the number omitted.
Parts: 2; items per part: 14; working time: 16 minutes.

CBT - Cognition of Behavioral Transformations

The ability to redefine behavioral information.

Picture Exchange CBT03A. Which alternative, when substituted for the picture indicated by the arrow, will change the meaning of the story?

Answer: 2. Score: number of items right plus one-third of the number omitted.
Parts: 2; items per part: 9/12, 9/12; working time: 16 minutes.

Social Translations CBT02A. Between which alternative pair will the given statement have a different intention or meaning?

parent to child

"I don't think so."

1) teacher to student
2) student to teacher
3) student to student

Answer: 2. Score: number of items right plus one-third of the number omitted.
Parts: 2; items per part: 11/12, 12/12; working time: 8 minutes.

CBI - The ability to predict what other people will do in given behavioral situations.

Cartoon Predictions CBI03A. Which alternative situation can be predicted from the given one?

Answer: 1. Score: number of items right plus one-third of the number omitted.
Parts: 2; items per part: 15/15, 14/15; working time: 8 minutes.

Factors of Social Creativity and their Tests

DBU - Divergent Production of Behavioral Units

The ability to create rapidly and fluently many responses that fit certain behavioral specifications.

Alternate Picture Meanings - DBU02A. Write many different things that a person might say if he felt as the person in a given picture does.

Sample Item:

Score: Number of feelings or thoughts judged to be behaviorally distinct.

Parts: 4; items per part: 2; working time: 12 minutes.

1. Let me see. Where was I?
2. I wish he'd shut up!
3. I can't study anymore tonight.
4. Good Grief! What have I done?
5. Why doesn't he leave me alone?

Expressing Mixed Emotions - DBU04A. Write many different things that a person might say when he is feeling both of two given emotions.

Sample Item: both JEALOUS and DISAPPOINTED.

1. You can have it. I don't want it anyway.
2. Yeah, Bill won; he always wins.
3. He got it? But I expected to.

Score: Number of behaviorally different responses indicating that both emotions are being experienced.

Parts: 4; items per part: 2; working time: 15 minutes.

Alternate Social Meanings - DBU01A. given an action of a person, write many different interpretations, each showing how the person might think or feel.

Sample Item: If one person winks at another, what could he (she) be thinking or feeling?

1. How about a date?
2. Here he comes now.
3. The person is being friendly.
4. You and I know better.
5. You're cute.
6. The person is kidding.
7. Don't give the secret away.

Score: Number of thoughts or feelings judged to be behaviorally different.
Parts: 4; items per part: 1; working time: 8 minutes.

DBC - Divergent Production of Behavioral Classes

The ability to classify and reclassify behavioral information into different categories of behavioral meaning.

Alternate Expressional Groups - DBC03A. Group given expressions in many different ways so that each group of at least 3 pictures expresses a different thought, feeling, attitude, or intention.

Sample Item:

Score: Number of appropriate groups produced that are judged to be behaviorally different.

Group 1 A,B,F Group 2 A,C,E

Parts: 2; items per part: 1; working time: 6 minutes.

Multiple Behavioral Grouping - DBC02A. Group given comments into many different sets according to the thoughts, feelings, or intentions they express.

Sample Item:

1. You get out of here
2. Are you sure
3. What a bore
4. How could you do such a thing
5. Didn't you listen to me
6. I wonder what time it is

Group A 1 3 4 5

Group B 2 4 5

Group C

Score: Number of appropriate groups of three or more items that indicate different behavioral classes.

Parts: 4; items per part: 1; working time: 8 minutes.

DBR - Divergent Production of Behavioral Relations

The ability to create many behaviorally meaningful relationships among people who are pictured or described, the relationships being appropriate to the expressions or situations of the people.

Alternative Facial Relations - DBR01A. Given photographs of different facial expressions and a comment, choose many different pairs of faces such that the first face chosen in each pair is making the comment to the second one.

Sample Item:

Comment: "Wait, that's not what I really meant.

Score: Number of appropriate pairs judged to be behaviorally different.

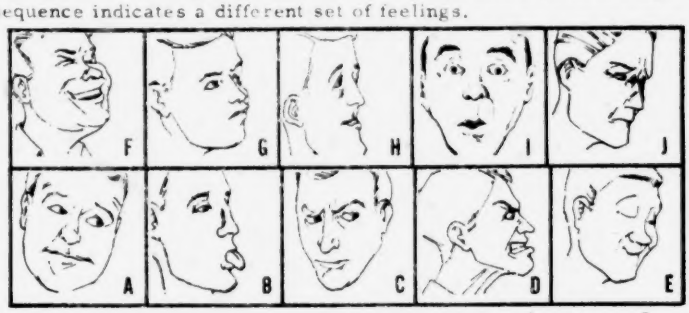
Parts: 3; items per part: 3; working time: 9 minutes.

Relation 1: B and C Relation 2: ___ and ___ Relation 3: ___ and ___

Multiple Expression Changes - DBT02A. Choose many different sequences of faces that show how a person might feel at different points of a given story, so that each sequence indicates a different set of feelings.

Sample Item:

A man trips a lady who is walking by.
She falls, and the man then apologizes to her.
The lady then becomes angry.



	1	2
Show how the man might feel as he trips the lady:	I	
Show how the man might feel as he apologizes to her:	A	
Show how the man might feel after the lady becomes angry:	C	

Score: Number of appropriate sequences that are judged behaviorally different.

Parts: 4; items per part: 1; working time: 12 minutes.

Varied Emotional Relations - DBR03A. From many pictures of individuals, choose many different sets of two pictures, each set showing a cause-effect relationship.

Sample Item:

Relation 1. C and A Relation 2. _____ and _____ Relation 3. _____ and _____

Score: Number of appropriate pairs showing behaviorally different relationships.
Parts: 2; items per part: 1; working time: 4 minutes.

DBS - Divergent Production of Behavioral Systems

The ability to produce fluently appropriate complex behavioral interrelationships among several people who are pictured or described as interactors.

Writing Behavioral Stories - DBS03A. Given a photograph of three people in a social situation, write many different stories describing how the people feel, and what they are thinking, and why.

Sample Item:

1. The blond girl, B, feels sorry for A because she thinks he is sick. C thinks he's faking, but A likes the attention
2. B tells A she's sorry she can't go to the dance with him. C feels sorry for A, but A is only faking unhappiness because he really wants to ask C.

Score: Number of behaviorally different stories interrelating the feelings and attitudes of the three people.

Parts: 3; items per part: 1; working time: 12 minutes.

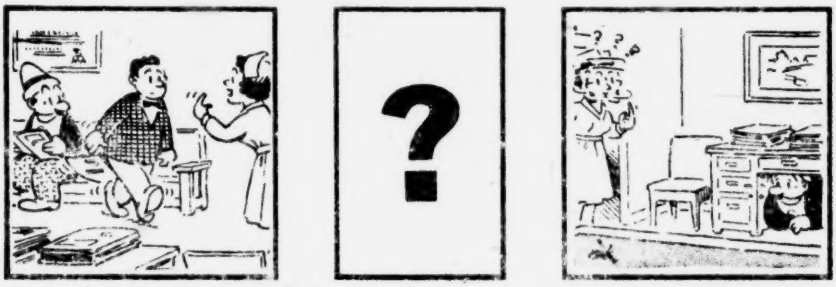
42. Sex. Females were assigned a value of 0; males a value of 1.

DBT - Divergent Production of Behavioral Transformations

The ability to reinterpret flexibly social situations in different behavioral meanings.

Multiple Cartoon Fill-Ins - DBS02A. Given the first and last frames of a cartoon strip, write what might have happened between them so that the explanation involves the feelings, thoughts, and intentions of the cartoon characters.

Sample Item:



1. He heard the man holler with pain and became afraid.
2. He thought the nurse was cute and played hide and seek.

Score: Number of behaviorally different responses that interrelate the characters and account for the behavior in both the first and last frames.
Parts: 3 (Note: only parts I, III, & IV were used); items per part: 1; working time: 9 minutes.

DBI - Divergent Production of Behavioral Implications

The ability to anticipate many different consequences or interpretations of interpersonal events.

Behavioral Elaboration - DBI04A. Given an action of one person, write many different responses to show how a second person might feel or react.

Sample Item:

IF PERSON A WINKS AT PERSON B, WHAT WILL B DO?

1. Smile back timidly
2. Look surprised
3. get embarrassed and blush
4. pretend she doesn't see person A
5. get mad at A

Score: Number of behaviorally different responses that are expectable from the given situation.

Parts: 4; items per part: 1; working time: 12 minutes.

Multiple Social Problems - DBI03A. Given two members of a typical family, write many different personal problems that they might have with each other. The problems should involve the feelings, thoughts, and attitudes of the two given people.

Sample Item:

What personal problems can the BROTHER and SISTER have with each other?

1. Sister makes fun of brother's friends
2. Brother and sister compete for attention of mother
3. Brother tries to dominate younger sister

Score: Number of interpersonal problems judged to be behaviorally different.

Parts: 4; items per part: 1; working time: 8 minutes.

This publication is published pursuant to a contract with the U.S. Office of Education, Department of Health, Education and Welfare. Points of view or opinions stated do not necessarily represent official U.S.O.E. position or policy.